D.V. Kolosov, D.A. Kuznetsova
Ural Federal University named after the first President of Russia B.N. Yeltsin
Yekaterinburg, Russia

DATA SCIENCE IN COMMUNICATION SPHERE

Abstract: The article describes application of the data science sphere in communication processes. It covers advertising technology, artificial intelligence, voice assistants, automated chat bots. Moreover, the article provides an opinion about the prospects of data science in our everyday life and communication between people, computers and the whole society.

Keywords: data science, artificial intelligence, computer science, natural language processing, deep/machine learning.

Д.В. Колосов, Д.А. Кузнецова
Уральский федеральный университет имени первого Президента России Б.Н. Ельцина
Екатеринбург, Россия

НАУКА О ДАННЫХ В СФЕРЕ КОММУНИКАЦИЙ

Аннотация: Статья описывает применение науки о данных в процессах коммуникации. Рассматриваются технологии рекламы, искусственный интеллект, голосовые помощники и автоматизированные чат боты. Кроме того, уделяется внимание перспективам применения науки о данных в повседневной жизни и коммуникации между людьми и компьютерами.

Ключевые слова: наука о данных, искусственный интеллект, компьютерная наука, обработка естественного языка, машинное обучение, нейронные сети.

The evolution of computer systems is constant. The past two decades have witnessed enormous leaps in the development of computers and the Internet. Commercial and commonly used Internet, which appeared in the early 2000s, has taken communication to a new level. This period changed the life of modern humanity and reduced the influence of communicational barriers in most of the spheres.

Communication and, generally, the way people talk have increased in the amount, variance and efficiency since the rise of the modern Internet and the consequences are inevitable. The Internet, if simplified, is a huge distributed storage of different data which is exchanged between participants of the network. This means that this data is stored somewhere and can be managed. People generate a huge amount of data every second. Messages, calls, clicks and just the fact of a person viewing a page are recorded. Business organized our communication on the Internet and business has a well-defined rule about the possessions: «Never lose an opportunity». In this case, the opportunity to process and analyze data created by engineers has escalated to a wide branch of study – *data science*.

Data science and the whole IT have deeply infiltrated and extended the communication from all the sides of this concept. This article describes the most common and perspective fields of application that people either already experience or will experience in the nearest future (Figure 1) [1].

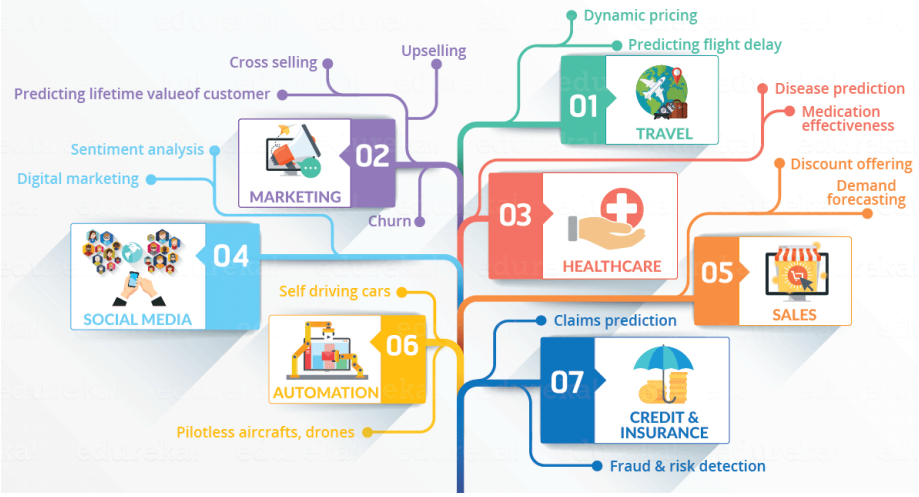

Figure 1 – Data science applications

The field includes three main parts:

- Collecting and storing data
- Processing data
- Interpreting analysis results

The latter two parts are of the highest interest. Apart from the fact that corporations know almost everything about us just because the Internet is a major part of our lives and use this knowledge to make customer-oriented research, fight cyber-crime and generally hold the information monopoly, engineers are developing digital products that are now perceived as extremely futuristic.

The most common sphere, where data science is applied, is advertising. People are exposed to a large number of advertisements hourly and most of them are intersecting with their interests at least partly. Thus, gamers get ads about new games, keen travellers get accommodation and tour advertisements, programmers and marketers get professional qualification courses or services. The ads provided to us are carefully generated by the data analysis algorithms which take every click, «like», time spent on the web page and other parameters into consideration to form our customers profile and a unique set of personal traits. The data collected from us is compared to the advertising preferences and a *targeted ad* is made.

Promoted posts on Instagram, YouTube ads, Facebook banners, Google and many more tools are leading and pushing us to buy something or at least attracting attention to the subject of promotion (Figure 2) [4].

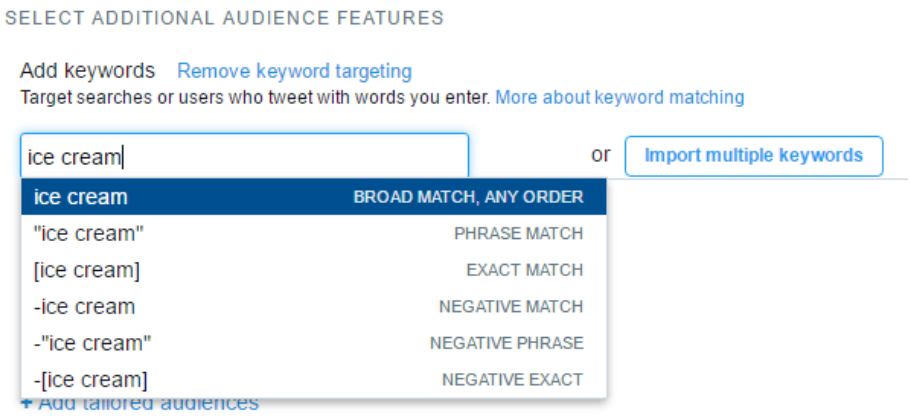

Figure 2 – Twitter Advertising Platform

Major analysis field implemented into such technology is called artificial intelligence – a statistical approach that consumes data in order to find patterns in it, especially in people's behaviour. In case of advertising, such software is responsible for pointing out that people with one set of

traits find a particular advertisement interesting more often than others. The statistics is able to give the most powerful insights which no algorithm or formula even come close to provide.

However, the use of the data science is more about utility rather than direct work with information. Scientists and engineers all over the world struggle to solve another big problem – making a computer understand the human language in terms of speech and text. This formed another sphere of data science which is called NLP or Natural Language Processing.

Why is it so important? A significant part of global workforce is wasted on dull communicative jobs such as customer relations, support agents, receptionists, consultants in facilities such as airports or railway stations and so on. Money and time spent on both of the sides (business and customer) could be reduced by replacing these jobs with automated software agents capable of doing the same function even more efficiently than a human.

NLP is constantly developing. The problem of recognizing speech is close to being fully solved. Voice assistants in mobile phones and homes (at least in case with English) are already capable of clearly understanding what is said in most of the cases. Language analysis part is not far behind – the technology tracks and selects relevant entities in your sentences in order to process them and generate the necessary output in the form of a sentence or action. The examples are placed in most of the modern devices: Siri from Apple, Google Assistant or Cortana from Microsoft are complex systems with plenty of knowledge in their memory. What is more, they are constantly learning because millions of people speak to them, leave feedback and simply make them work to get more data and more «knowledge» (Figure 3) [6].

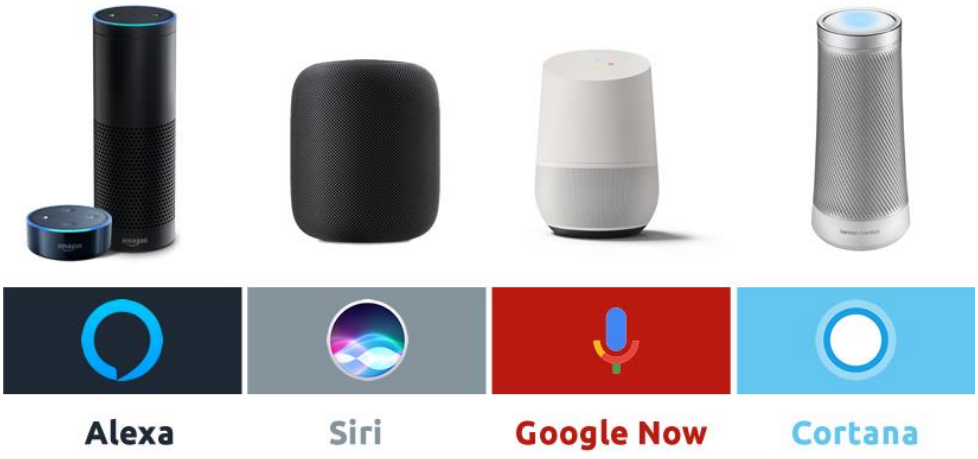

Figure 3 – Voice assistants

An example of the system under the hatch of such products is DialogFlow by Google (Figure 4). This product is capable of selecting relevant words out of the text and building triggers based on the words, combination or the whole sentences. DialogFlow is frequently used as a basis for different programs interacting with human natural language. [2]

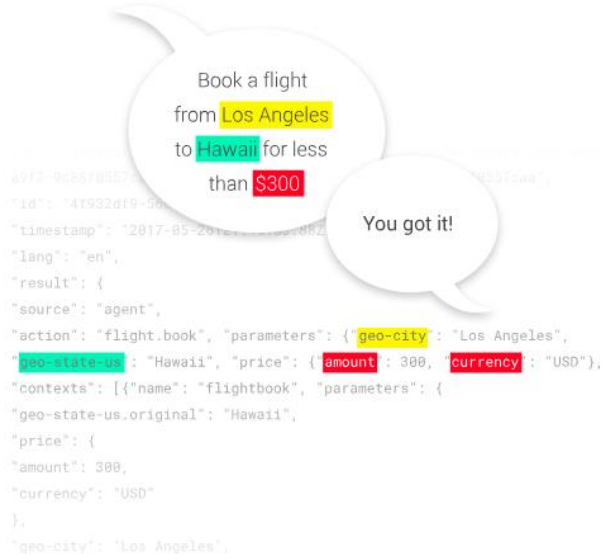

Figure 4 – DialogFlow demonstration

With the amount of data collected every second and the amount of effort put into pushing the borders of artificial intelligence, data science and NLP, machines will soon be able to act like people and fully understand them.

The controversy of the question of artificial intelligence implementation is rising because people do not trust technology, do not want to lose their jobs or they are just scared that the AI can become the next step in evolution and overcome Homo Sapiens dominance on the planet. There is always a risk, but with extreme growth of population, the more social automation we can create, the more time we will have to solve more crucial tasks like saving the whole planet and humankind from the threats that keep on pushing.

Data science can bring high efficiency and convenience to our communication with each other. Information is the new oil and soon it will be accepted globally

REFERENCES

1. Data science applications. [Electronic recource]. URL: https://www.houseofbots.com/ images/ news/4386/cover.png (29.01.19)
2. DialogFlow main page. [Electronic recource]. URL: https://dialogflow.com/ (29.01.19).
3. DialogFlow demonstration. [Electronic recource]. URL: https://dialogflow.com/images/ api_home_code.svg (29.01.19).
4. Twitter advertising page: Audience targeting. [Electronic recource]. URL: https://marketing.twitter.com/na/en/solutions/create-engagement/audience-targeting.html (29.01.19).
5. Twitter Advertising Platform. [Electronic recource]. URL: https://www.socialmedia examiner.com/wp-content/uploads /2017/03/ab-twitter-ad-targeting-by-keyword.png (29.01.19).
6. Voice assistants: How artificial intelligence assistants are changing our lives every day. [Electronic recource]. URL: https://www.smartsheet.com/voice-assistants-artificial-intelligence (31.01.19).
7. Voice assistants. [Electronic recource]. URL: https://geeksfl.com/wp-content/uploads/ voice-assistants-battle.jpg (31.01.19).